Submitted to Journal of The Acoustical Society of America

3P

NOTE ON BINAURAL MASKING LEVEL DIFFERENCES AT HIGH FREQUENCIES*

$2.75 P.

N. I. Durlach

Document

Center for Communications Sciences, Research Laboratory of Electronics

Massachusetts Institute of Technology

Cambridge, Massachusetts

FACILITY FORM 602

N64-27809 (ACCESSION NUMBER)

27 (PAGES)

CR-56720 (NASA CR OR TMX OR AD NUMBER)

(THRU) 1

(CODE) 08

(CATEGORY)

Abstract

27809

In this note, a quantitative "black-box" model is developed for use in interpreting certain data on binaural masking level differences at high frequencies. The basic idea of this model is that these differences are the result of variations in the extent to which the envelopes of the signals presented to the two ears of the listener are unequal.

AUTHOR

OTS PRICE

XEROX $ 2.00 ph

MICROFILM $ _____

* This work was supported in part by the U. S. Army Signal Corps, the Air Force Office of Scientific Research, and the Office of Naval Research; in part by the National Science Foundation (Grant G-16526) and the National Institutes of Health (Grant MH-04737-03); and in part by the National Aeronautics and Space Administration (Grant NsG-496).

The purpose of this note is to put in quantitative form a model of binaural masking level differences described qualitatively in the writer's recent paper on the Equalization and Cancellation (EC) Model.[1] The introduction of this model is motivated by the inapplicability of both the EC model and the Phase Detector (PHD) Model[2,3] to the experimental results at high frequencies. In the present model, as opposed to the two just cited, it is assumed that the phase information is lost at the periphery and that the binaural interaction is based purely on intensity differences. Although this model can be regarded as a high frequency adjunct to either of the other two models, in its present form, it is a more natural extension of the PHD model than of the EC model. Specifically, it is nothing more than the amplitude-domain counterpart of the PHD model.[4]

Inasmuch as there exists a substantial amount of data that is consistent with the notion that the auditory system does not make use of phase information in high-frequency signals,[5] the general idea that masking phenomena at these frequencies are based on amplitude or intensity comparisons is regarded by the writer as a reasonable one. However, the way in which this idea is incorporated into a specific quantitative model is rather arbitrary. As in the writer's previous work, the auditory system is treated as a "black box" and no account is taken of the peripheral coding into nerve impulses or of the biological data on the centers of binaural interaction. In order to make the model meaningful, one must assume that it can be translated into a physiological model in such a way that the quantities of interest remain

invariant under the translation.

As in the writer's paper on the EC model, attention is restricted to stimuli in which (a) the signal to be detected is a pulsed tone of relatively long duration (duration $\geq$ 0.1 sec); (b) the masking signal is white Gaussian random noise of relatively large bandwidth (large enough to contain the critical band); (c) the power levels of the signals are relatively high (noise spectrum level $\geq$ 40 dB re 0.0002 dyn/cm^2); (d) the stimuli are presented by means of earphones (the interaural relations being controlled by means of electronic circuitry). The data to be considered arise from the work of Hirsh,[6] Hirsh and Webster,[7] Hirsh and Burgeat,[8] Webster,[2] Jeffress, Blodgett, and Deatherage,[9,10] Jeffress, Blodgett, Sandel and Wood,[3] Blodgett, Jeffress, and Taylor,[11] Green,[12] and Durlach.[1] These sources will be denoted H, HW, HB, W, JBD, JBSW, BJT, G, and D, respectively. Inasmuch as the model is appropriate only to frequencies at which the phase information is not used, attention will be focused on frequencies above approximately 1400 cps; however, in order to put the results in a proper setting, data will also be plotted for lower frequencies.

Let

$s(t)$ = signal tone

$n(t)$ = masking noise

$y_j(t)$ = total signal to ear j ($j = 1, 2$).

The seven types of stimuli to be considered (denoted $E_{m,m}$, $E_{o,o}$, $E_{\pi,\pi}$, $E_{m,o}$, $E_{m,\pi}$, $E_{\pi,o}$, and $E_{o,\pi}$) are defined in Table I. For any stimulus E among these seven, let r(E) denote the input

signal-power-to-noise-power ratio corresponding to the detection threshold. Roughly speaking, the data on the thresholds for these stimuli for frequencies above ≈ 1400 cps indicate the following relations:

$$r(E_{m,m}) = r(E_{o,o}) = r(E_{\pi,\pi}) = r(E_{m,o}) = r(E_{m,\pi})$$
$$= 2r(E_{\pi,o}) = 2r(E_{o,\pi}).$$

Of these results, the chief surprise, according to those concerned with this field, is the lowered threshold for the stimuli $E_{\pi,o}$ and $E_{o,\pi}$. A frequently voiced opinion has been that, "since the auditory system is insensitive to phase at these frequencies, changing the interaural phase of the signal or noise component should have no effect on the threshold." That the conclusion of this argument is incorrect is proved by the data. The nature of the error in this argument, however, depends upon the meaning one assigns to the expression "insensitive to phase." If one means that that the auditory system is insensitive to changes in the relative phases of the components of a stimulus, then the conclusion is based on a false premise.[13] If, on the other hand, one means that the auditory system is insensitive to changes in the phase of the total stimulus (i.e., only the envelope of the total stimulus is retained), then, although the premise may be true, the conclusion does not follow from it. In this case, the premise is relevant only to those binaural masking stimuli in which the interaural phase shift of the signal component is identical to that of the noise component. As will be seen below, changes in the interaural phase of one component relative to the other component produce corresponding changes in the interaural envelope relations.

Thus, even if the auditory system is capable only of observing the envelopes of the stimuli, there is no reason to believe that the threshold should remain fixed when the phase of one of the components is changed.

Denoting the angular frequency of the tone $s(t)$ by ω and assuming that the noise $n(t)$ has been passed through a bandpass filter centered on ω (assumed to correspond to the critical band around ω), one can write

$$s(t) = a_s \cos(\omega t) \tag{1}$$

$$n(t) = a_n(t) \cos\left[\omega t + \phi(t)\right] \tag{2}$$

$$y_j(t) = v_j(t) \cos\left[\omega t + \psi_j(t)\right] , \tag{3}$$

where the amplitudes $a_n(t)$ and $v_j(t)$ and phases $\phi(t)$ and $\psi_j(t)$ are random functions that are slowly varying in comparison with the carrier function $\cos(\omega t)$.

Assume now that the binaural processing system has available to it only the power functions $v_1^2(t)$ and $v_2^2(t)$, the phase functions $\psi_1(t)$ and $\psi_2(t)$ being lost at the periphery.[14] Inasmuch as the stimuli $E_{\pi,\pi}$, $E_{m,\pi}$, and $E_{o,\pi}$ differ from the stimuli $E_{o,o}$, $E_{m,o}$, and $E_{\pi,o}$ (respectively) only in the sign of $y_2(t)$, one concludes immediately that

$$r(E_{\pi,\pi}) \,/\, r(E_{o,o}) = 1 \tag{4}$$

$$r(E_{m,\pi}) \,/\, r(E_{m,o}) = 1 \tag{5}$$

$$r(E_{o,\pi}) \,/\, r(E_{\pi,o}) = 1. \tag{6}$$

A comparison of these equations with the experimental data is shown in Figs. 1, 2 and 3. (Note that in all figures presented

in this paper the masking level differences are plotted in dB.) Roughly speaking, whereas the first equation is consistent with the data for all frequencies tested, the second two are consistent with the data for all frequencies tested above $\approx$ 400 cps.[15] Since only the high frequencies are of interest here, in the following discussion, attention will be restricted to the stimuli $E_{m,m}$, $E_{o,o}$, $E_{m,o}$, and $E_{\pi,o}$.

Assume, furthermore, that the quantity of concern in the binaural detection procedure is the interaural power-difference function

$$P(t) = v_1^2(t) - v_2^2(t). \tag{7}$$

This quantity plays the same role in the present model as the interaural phase-difference function plays in the PHD model. The total detection procedure is assumed to be a combination of the monaural detection procedure (a process that operates on each of the terms $v_j^2(t)$ taken by itself) and a binaural detection procedure that operates on $P(t)$. Inasmuch as $P(t)$ is always identically zero for the stimulus $E_{o,o}$ (independent of whether or not the signal to be detected is present), one concludes that there will be no binaural improvement for $E_{o,o}$ and that

$$r(E_{o,o}) \,/\, r(E_{m,m}) = 1. \tag{8}$$

A comparison of this equation with the data is shown in Fig. 4. As in the case of the ratio $r(E_{\pi,\pi}) \,/\, r(E_{o,o})$, theory and experiment appear to be consistent for all frequencies tested. In the remarks that follow, no distinction will be made between the

stimuli $E_{o,o}$ and $E_{m,m}$.

For the stimuli $E_{\pi,o}$ and $E_{m,o}$ the quantity P(t) is given[16] by

$$E_{\pi,o}: \quad P(t) = 4a_s a_n(t)\cos\left[\phi(t)\right] \tag{9}$$

$$E_{m,o}: \quad P(t) = a_s^2 + 2a_s a_n(t)\cos\left[\phi(t)\right] . \tag{10}$$

When noise alone is present, P(t) is identically zero for both stimuli. Thus the occurrence of a nonzero value for P(t) implies the presence of a signal. In order to incorporate this fact into a detection model, it will be assumed that binaural detection occurs when some statistic Q related to P(t) exceeds a certain binaural threshold T_b. It will also be assumed that there is no interaction between monaural detection and binaural detection and that the ultimate decision as to the presence or absence of a signal is the logical sum of the two component decisions. In other words, it will be assumed that the total detection procedure results in the output "signal present" if and only if the monaural detection procedure taken alone results in this output or the binaural detection procedure taken alone results in this output.[17] The threshold T_b will be assumed to be independent of frequency and independent of the type of stimulus. (This threshold plays the same role in the present model as the "fixed, minimum, interaural time delay" plays in the PHD model.) Finally, it will be assumed that the time interval during which the stimulus is observed is sufficiently long with respect to the correlation time of n(t) to allow one to identify time overages over this interval with ensemble averages (denoted $\langle\ \rangle$).

The precise choice of the statistic Q is rather arbitrary; however, it is clear that neither the mean value of P(t) nor the maximum value of $|P(t)|$ is adequate. A choice of the form $Q = f(\langle P \rangle)$, where f is an arbitrary function, is inadequate because it fails to explain the result $r(E_{\pi,o}) = r(E_{m,m}) / 2$. Inasmuch as $\langle P \rangle = 0$ for $E_{\pi,o}$ (independent of the value of a_s), choosing $Q = f(\langle P \rangle)$ would imply that $r(E_{\pi,o}) = r(E_{m,m})$. Going to the other extreme, it is equally clear that choosing $Q = \text{Max} |P(t)|$ is inadequate since, given a long enough time interval, $|P(t)|$ will exceed any given threshold T_b, provided only that $a_s \neq 0$. Thus, the choice of this statistic would imply that the detection performance could be made as good as one pleases merely by extending the interval during which the stimulus is observed. With regard to statistics of the form $Q = f(\langle P \rangle)$, it should also be noted that the rejection of these statistics is equivalent to the assumption that the binaural auditory system is sensitive to the fluctuations in P(t). Inasmuch as the correlation time of these fluctuations is the reciprocal of the critical bandwidth (of the order of a few milliseconds), this assumption does not appear unreasonable.[18]

The two choices for Q that will be considered are

$$Q_\alpha = \langle P^2 \rangle^{1/2} \tag{11}$$

$$Q_\beta = \langle P \rangle + \left[2(\langle P^2 \rangle - \langle P \rangle^2)\right]^{1/2}. \tag{12}$$

The statistic Q_α is merely the rms value of P(t). The results of applying the statistic Q_β to Eqs. (9) and (10) can be obtained directly from these equations by replacing $a_n(t)$ with $\langle a^2 \rangle^{1/2}$ and replacing $\phi(t)$ with a value of ϕ that produces a maximum,

i.e., $\phi = 0$. (This procedure is analogous to the one used by Webster for the PHD model.[2]) Denoting the statistic Q evaluated for the stimulus E by Q(E), one obtains

$$Q_{\alpha}(E_{\pi,o}) = 2\sqrt{2}\, a_s \langle a_n^2 \rangle^{1/2} \tag{13}$$

$$Q_{\alpha}(E_{m,o}) = (a_s^4 + 2a_s^2 \langle a_n^2 \rangle)^{1/2} \tag{14}$$

$$Q_{\beta}(E_{\pi,o}) = 4\, a_s \langle a_n^2 \rangle^{1/2} \tag{15}$$

$$Q_{\beta}(E_{m,o}) = a_s^2 + 2a_s \langle a_n^2 \rangle^{1/2}. \tag{16}$$

For any given Q and E, binaural detection is assumed to occur when $Q(E) \geq T_b$.

Assuming that E is a stimulus in which the detection of the signal is effected binaurally (as opposed to monaurally), one can compute the value of T_b (as a function of $\langle a_n^2 \rangle$) by specifying the signal-to-noise ratio $a_s^2 / \langle a_n^2 \rangle$ at threshold and equating Q(E) with T_b. According to the data shown in Figs. 5 and 6 for frequencies above $\approx$ 1400 cps, one has, approximately,

$$r(E_{o,o}) / r(E_{\pi,o}) = 2 \tag{17}$$

$$r(E_{o,o}) / r(E_{m,o}) = 1. \tag{18}$$

The fact that the first ratio is greater than unity implies that the detection of the signal in the stimulus $E_{\pi,o}$ is achieved binaurally. The fact that the second ratio is equal to unity implies either that the detection of the signal in the stimulus $E_{m,o}$ is achieved monaurally or that the monaural procedure and binaural procedure lead to detection results that are identical. Letting K denote the signal-to-noise ratio $a_s^2 / \langle a_n^2 \rangle$ at threshold

for the stimulus $E_{o,o}$, one can restate the results described by Eqs. (17) and (18) in terms of the threshold values[19] of $a_s^2 / \langle a_n^2 \rangle$ by writing

$$E_{o,o}: \quad a_s^2 / \langle a_n^2 \rangle = K \tag{19}$$

$$E_{\pi,o}: \quad a_s^2 / \langle a_n^2 \rangle = K/2 \tag{20}$$

$$E_{m,o}: \quad a_s^2 / \langle a_n^2 \rangle = K \, . \tag{21}$$

Assume now that the monaural detection procedure consists of comparing $\langle v_j^2 \rangle$ with a monaural threshold T_m. For the seven stimuli considered in this paper, it can be assumed without loss of generality that the monaural procedure operates on ear 1 only. Thus, monaural detection occurs when $\langle v_1^2 \rangle = a_s^2 + \langle a_n^2 \rangle \geq T_m$. Making use of Eqs. (13) - (16) and (19) - (21), and denoting the binaural threshold T_b derived from the statistic Q and stimulus E by $T_b(Q,E)$, one obtains for the various thresholds

$$T_m(K) = (1 + K) \langle a_n^2 \rangle \tag{22}$$

$$T_b(Q_\alpha, E_{\pi,o})(K) = 2\sqrt{K} \langle a_n^2 \rangle \tag{23}$$

$$T_b(Q_\alpha, E_{m,o})(K) = \sqrt{K^2 + 2K} \langle a_n^2 \rangle \tag{24}$$

$$T_b(Q_\beta, E_{\pi,o})(K) = 2\sqrt{2K} \langle a_n^2 \rangle \tag{25}$$

$$T_b(Q_\beta, E_{m,o})(K) = (K + 2\sqrt{K}) \langle a_n^2 \rangle . \tag{26}$$

According to these equations, all of the thresholds vary linearly with the noise power. In order to see how the thresholds depend upon the value chosen for K, the coefficients of $\langle a_n^2 \rangle$ are plotted in Fig. 7. [The crossover points $K_o(Q_\alpha)$, $K_o(Q_\beta)$, $K_o'(Q_\alpha)$, and $K_o'(Q_\beta)$ are discussed below.]

If the monaural threshold T_m is exceeded at a lower signal-to-noise ratio $a_s^2 / \langle a_n^2 \rangle$ than the binaural threshold T_b for the stimulus $E_{m,o}$ (so that the detection of the signal in $E_{m,o}$ is effected monaurally), then the values of T_b derived from the $E_{m,o}$ data are irrelevant. If, on the other hand, T_m is exceeded at the same value of $a_s^2 / \langle a_n^2 \rangle$ as T_b (so that the detection of the signal in $E_{m,o}$ can equally well be regarded as occurring binaurally), then these values are relevant and they must be consistent with those obtained from the stimulus $E_{\pi,o}$. Specifically, since the threshold T_b was assumed to be independent of the type of stimulus, one must have $T_b(Q,E_{m,o})(K) = T_b(Q,E_{\pi,o})(K)$. Denoting the solution of this equation by $K = K_o(Q)$, one obtains

$$K_o(Q_\alpha) = 2.00 \tag{27}$$

$$K_o(Q_\beta) = 0.69 \tag{28}$$

$$T_b(Q_\alpha,E_{m,o})\left[K_o(Q\ )\right] = T_b(Q_\alpha,E_{\pi,o})\left[K_o(Q_\alpha)\right] = 2.83\langle a_n^2\rangle \tag{29}$$

$$T_b(Q_\beta,E_{m,o})\left[K_o(Q_\beta)\right] = T_b(Q_\beta,E_{\pi,o})\left[K_o(Q_\beta)\right] = 2.35\langle a_n^2\rangle. \tag{30}$$

The constraint on K implied by the fact that T_b cannot be exceeded at a lower value of $a_s^2 / \langle a_n^2 \rangle$ than T_m for the stimulus $E_{m,o}$ is given by $K \leq K_o(Q)$. [The equality in this relation not only provides a solution to the equation $T_b(Q,E_{m,o})(K) = T_b(Q,E_{\pi,o})(K)$, but also specifies the value of K for which monaural detection and binaural detection occur at the same signal-to-noise ratio for the stimulus $E_{m,o}$.] For any choice of K that satisfies the relation $K \leq K_o(Q)$, the empirical Eqs. (9) and (10) can be derived from the model by choosing the binaural threshold to be $T_b(Q,E_{\pi,o})(K)$. Finally, if one requires that the monaural threshold T_m and the binaural

threshold T_b be equal (a hypothesis that would make the model more economical), then K must satisfy the equation $T_m(K) = T_b(Q,E_{\pi,o})(K)$. Denoting the solution of this equation by $K_o'(Q)$, one obtains

$$K_o'(Q_\alpha) = 1.00 \tag{31}$$

$$K_o'(Q_\beta) = 0.17 \tag{32}$$

$$T_m\left[K_o'(Q_\alpha)\right] = T_b(Q_\alpha,E_{\pi,o})\left[K_o'(Q_\alpha)\right] = 2.00\langle a_n^2\rangle \tag{33}$$

$$T_m\left[K_o'(Q_\beta)\right] = T_b(Q_\beta,E_{\pi,o})\left[K_o'(Q_\beta)\right] = 1.17\langle a_n^2\rangle\ . \tag{34}$$

In considering these results on the value of K and the corresponding values of T_m and T_b, three points should be noted. First, the result K = 1 is merely a restatement of Fletcher's hypothesis[20] (used by Webster in his work on the PHD model[2]). Second, the values obtained for K, T_m, and T_b are obviously sensitive not only to the choice of statistic, but also, to the precise manner in which the data are fitted. A modest variation either in Q or in Eqs.(17) and (18) will result in an appreciable variation in the values for K, T_m, and T_b. Thus, one should not take the precise values of these parameters as specified by the above equations too seriously. Only the orders of magnitude of these quantities have any significance. Third, unless K is chosen to be appreciably less than unity, the thresholds T_m and T_b determined by K are considerably larger than what one would expect from considering the results on just-noticeable differences (jnds) in intensity. For example, if one chooses K = 1 and uses the statistic Q_α, one has $T_m = T_b = 2\langle a_n^2\rangle$. This implies that the signal goes undetected unless its presence

results in a power level of the total stimulus that is at least 3 dB above the power level of the noise. The results of jnd experiments, on the other hand, indicate that the jnd intensity threshold lies in the interval 0.5 to 2.0 dB.[21] (A similar discrepancy has been observed in the PHD and EC models in the time domain.[22])

In conclusion, one should note that, although the set of stimuli considered in this paper was restricted to those involving interaural phase shifts of magnitude π (i.e., changes of sign), it is obvious that the model can be applied to stimuli in which the phase shifts take on other values as well. However, insofar as there are no high frequency data for these other stimuli and, even if there were, the magnitude of the binaural masking level differences would be extremely small, generalizing the model did not appear to be worthwhile.

As stated previously, this model provides a natural extension of the PHD model to high frequencies. Roughly speaking, it can be derived from that model merely by replacing the word "phase" with the word "power" wherever the word "phase" occurs. In order to adjoin it to the PHD model, one only need make the assumption that, whereas at low frequencies, the variable of concern is the interaural phase difference, at high frequencies, the variable of concern is the interaural power difference. Whether or not another high-frequency model can be constructed which provides an equivalently natural extension of the EC model remains to be seen.

The writer is indebted to the following people for useful discussions concerning the material presented in this note: P. R. Gray, J. L. Hall II, L. A. Jeffress, A. W. Mills, T. T. Sandel, W. M. Siebert, J. V. Tobias, F. A. Webster, and J. R. Welch.

Footnotes and References

1. N. I. Durlach, "The Equalization and Cancellation Model of Binaural Masking Level Differences," J. Acoust. Soc. Am. 35, 1206-1218 (1963).

2. F. A. Webster, "The Influence of Interaural Phase on Masked Thresholds. I. The Role of Interaural Time Deviation," J. Acoust. Soc. Am. 23, 452-462 (1951).

3. L. A. Jeffress, H. C. Blodgett, T. T. Sandel, and C. L. Wood III, "Masking of Tonal Signals," J. Acoust. Soc. Am. 28, 416-426 (1956).

4. A somewhat different high-frequency model (but one that is also based on intensity comparisons) was developed by F. A. Webster in 1951 (unpublished).

5. For psychoacoustic data, see, for example, A. W. Mills, "Lateralization of High-Frequency Tones," J. Acoust. Soc. Am. 32, 132-134 (1960). For neurophysiological data, see, for example, N. Y-S. Kiang, T. Watanabe, E. C. Thomas, and L. F. Clark, "Stimulus Coding in the Cat's Auditory Nerve; Preliminary Report," Ann. Otol. Rhinol. Laryngol. 71, 1009-1026 (1962); also monograph in preparation by same authors.

6. I. J. Hirsh, "The Influence of Interaural Phase on Interaural Summation and Inhibition," J. Acoust. Soc. Am. 20, 536-544 (1948).

7. I. J. Hirsh and F. A. Webster, "Some Determinants of Interaural Phase Effects," J. Acoust. Soc. Am. 21, 496-501 (1949).

8. I. J. Hirsh and M. Burgeat, "Binaural Effects in Remote Masking," J. Acoust. Soc. Am. 30, 827-832 (1958).

9. L. A. Jeffress, H. C. Blodgett, and B. H. Deatherage, "The Masking of Tones by White Noise as a Function of the Interaural Phases of Both Components. I. 500 cycles," J. Acoust. Soc. Am. 24, 523-527 (1952).

10. L. A. Jeffress, H. C. Blodgett, and B. H. Deatherage, "Masking and Interaural Phase. II. 167 cps," J. Acoust. Soc. Am. 34, 1124-1126 (1962).

11. H. C. Blodgett, L. A. Jeffress, and R. W. Taylor, "Relation of Masked Thresholds to Signal-Duration for Various Interaural Phase-Combinations," Am. J. Psychol. 66, 283-290 (1958).

12. Green's data were obtained directly from D. M. Green through private communication. For a brief description of Green's experiment, see N. I. Durlach, op. cit., Footnote 29.

13. In order to demonstrate the absurdity of this premise, suppose, for example, that the two components are identical, each having the complex spectrum $F(\omega)$. The premise then implies that the auditory system is insensitive to the value of ϕ in the stimulus $F(\omega) + F(\omega)\exp(i\phi)$. In particular, choosing $\phi = 0$ and $\phi = \pi$, one is led to the absurd result that the auditory system cannot distinguish between the stimuli $2F(\omega)$ and 0. Similarly, if the complex spectrum of the total stimulus is given by $F(\omega) = |F(\omega)| \exp\left[i\phi(\omega)\right]$, and one takes the components of the stimulus to be the Fourier components of $F(\omega)$, then the premise implies that the auditory system is insensitive to variations in the phase spectrum $\phi(\omega)$. To see that this implication is absurd, one need only recall the fact that the energy spectrum of the Dirac delta function is identical to the power spectrum of white noise.

14. Since $v_j(t)$ is slowly varying in comparison with $\cos(\omega t)$, one can interpret $v_j^2(t)$ as being twice the average power in $y_j(t)$, where the average is taken over a cycle of $\cos(\omega t)$.

15. The slight discrepancy between theory and experiment for the ratio $r(E_{0,\pi}) \,/\, r(E_{\pi,0})$ at frequencies above 400 cps is judged by the writer to be insignificant.

16. For both $E_{\pi,o}$ and $E_{m,o}$ one has $v_1^2(t) = a_s^2 + a_n^2(t) + 2a_s a_n(t)\cos\left[\phi(t)\right]$. For $E_{\pi,o}$, one has $v_2^2(t) = a_s^2 + a_n^2(t) - 2a_s a_n(t)\cos\left[\phi(t)\right]$. For $E_{m,o}$, one has $v_2^2(t) = a_n^2(t)$.

17. The possibility that the total detection procedure might result in the output "signal present" even though neither of the component procedures taken alone results in this output will be ignored. L. A. Jeffress, H. C. Blodgett, T. T. Sandel, and C. L. Wood III, op. cit., refer to this possibility as "monaural contamination."

18. If τ denotes the length of the observation interval and w the width (in radians) of the critical band, the assumptions on the time constants in the stimulus are given by $\tau >> 2\pi/w >> 2\pi/\omega$. If the first relation is violated, one cannot replace time averages by ensemble averages. If the second relation is violated, then $v_j(t)$ will not be varying slowly in comparison with $\cos(\omega t)$ and the assumption that the auditory system is sensitive to fluctuations in P(t) will be inconsistent with the assumption that the auditory system cannot follow the carrier function $\cos(\omega t)$ and make use of the phase information in $\psi_j(t)$.

19. The only difference between the signal-to-noise ratio $a_s^2 / \langle a_n^2 \rangle$ at threshold and the signal-to-noise ratio r is that caused by the filtering action of the critical band. It is assumed implicityly throughout this paper that the critical band is independent of the type of stimulus.

20. For a comparatively recent discussion of Fletcher's hypothesis, see, for example, J. A. Swets, D. M. Green, and W. P. Tanner, Jr., "On the Width of Critical Bands", J. Acoust. Soc. Am. 34, 108-113 (1962).

21. See, for example, A. W. Mills, "Lateralization of High Frequency Tones", J. Acoust. Soc. Am. 32, 132-134 (1960); also further unpublished work by Mills.

22. In both the EC model and the PHD model, the time constant that results from applying the model to the data is considerably larger than the jnd in time (see N. I. Durlach, op. cit., p. 1218).

Figure Captions

1. Masking level difference $r(E_{\pi,\pi}) / r(E_{o,o})$ as a function of frequency. (Data obtained from H and HB.)

2. Masking level difference $r(E_{m,\pi}) / r(E_{m,o})$ as a function of frequency. (Data obtained from H and HB.)

3. Masking level difference $r(E_{o,\pi}) / r(E_{\pi,o})$ as a function of frequency. (Data obtained from H and HB.)

4. Masking level difference $r(E_{o,o}) / r(E_{m,m})$ as a function of frequency. (Data obtained from HB.)

5. Masking level difference $r(E_{o,o}) / r(E_{\pi,o})$ as a function of frequency. (Data obtained from H, HB, W, HW, JBD, BJT, JBSW, G, and D.)

6. Masking level difference $r(E_{o,o}) / r(E_{m,o})$ as a function of frequency. (Data obtained from H, HB, BJT, and G.)

7. Dependence of normalized threshold $T/\langle a_n^2 \rangle$ on the value of K.

Table I. Definition of Stimuli.[a]

Stimulus	$y_1(t)$	$y_2(t)$
$E_{m,m}$	$s(t) + n(t)$	0
$E_{o,o}$	$s(t) + n(t)$	$s(t) + n(t)$
$E_{\pi,\pi}$	$s(t) + n(t)$	$-s(t) - n(t)$
$E_{m,o}$	$s(t) + n(t)$	$n(t)$
$E_{m,\pi}$	$s(t) + n(t)$	$-n(t)$
$E_{\pi,o}$	$s(t) + n(t)$	$-s(t) + n(t)$
$E_{o,\pi}$	$s(t) + n(t)$	$s(t) - n(t)$

[a]The first subscript on E refers to the signal components and the second to the noise components. The subscript o denotes no interaural phase shift, the subscript π denotes an interaural phase shift of magnitude π, and the subscript m means that the given component was presented monaurally.

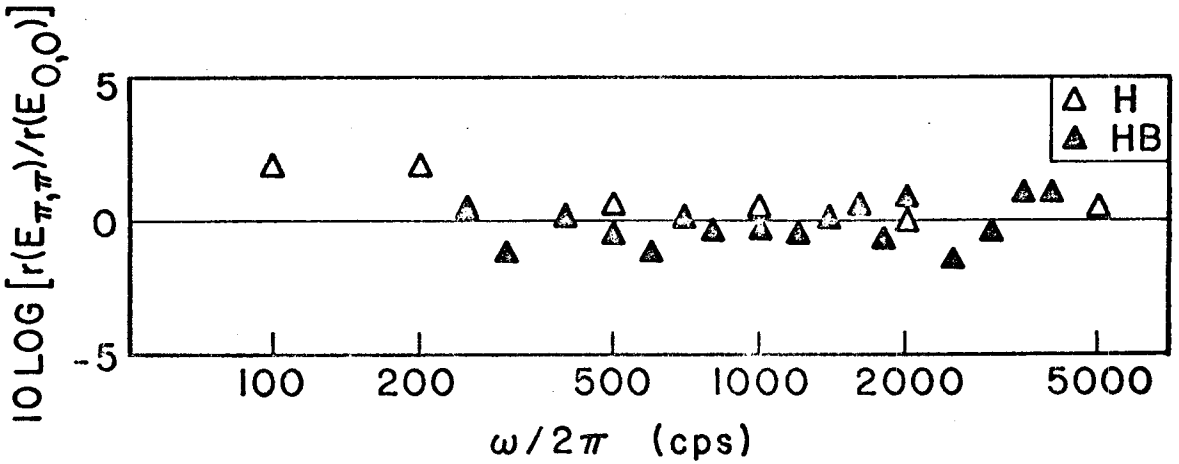

Fig. 1

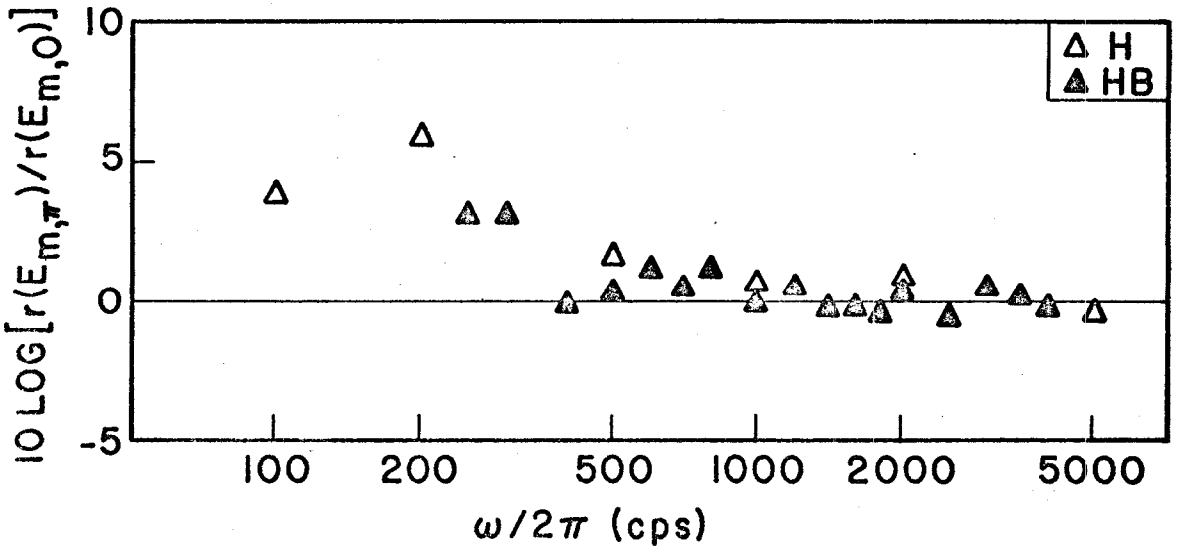

Fig. 2

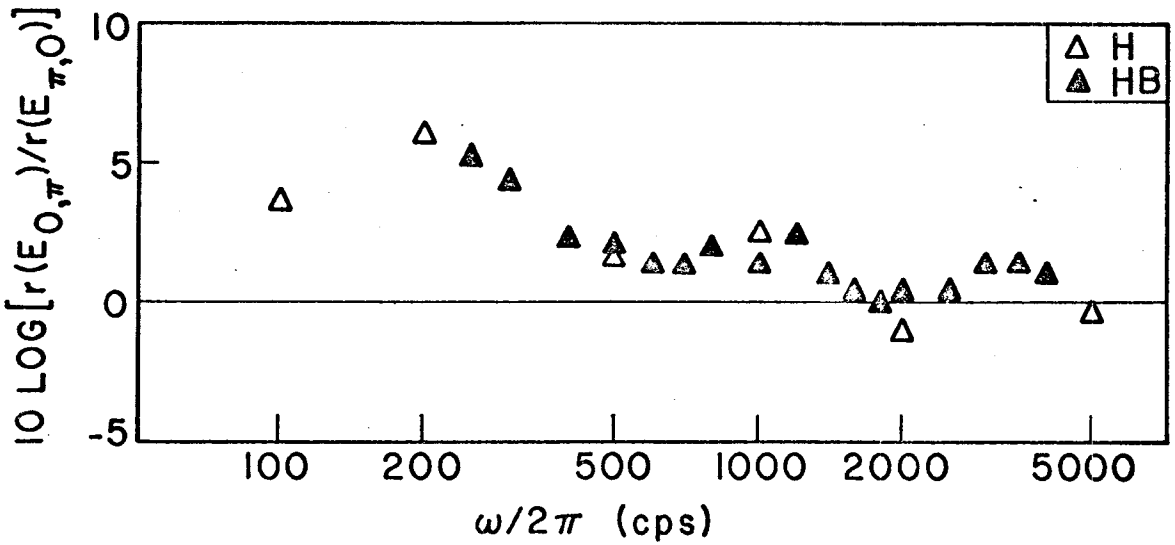

Fig. 3

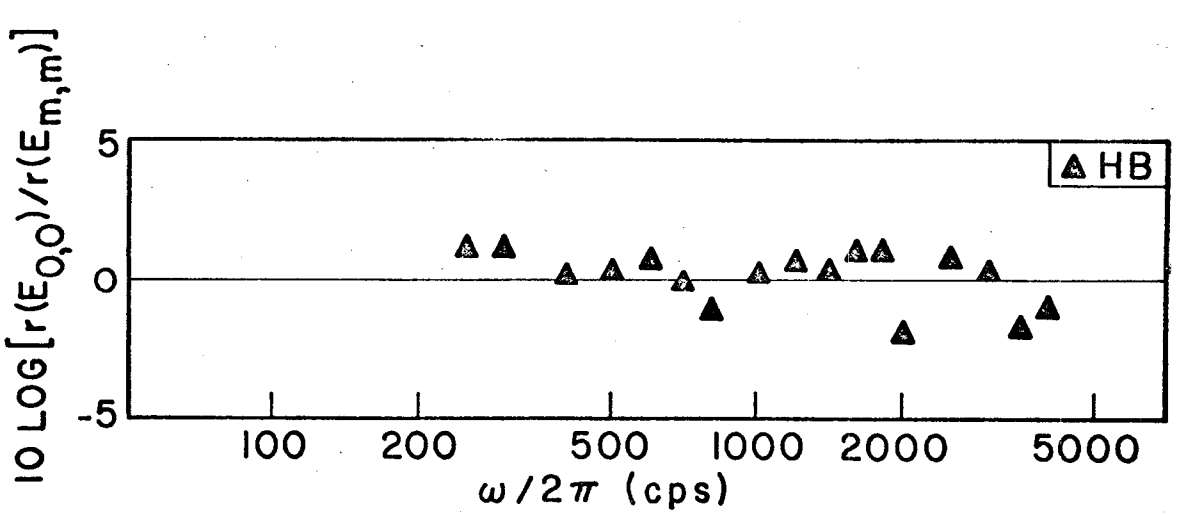

Fig. 4

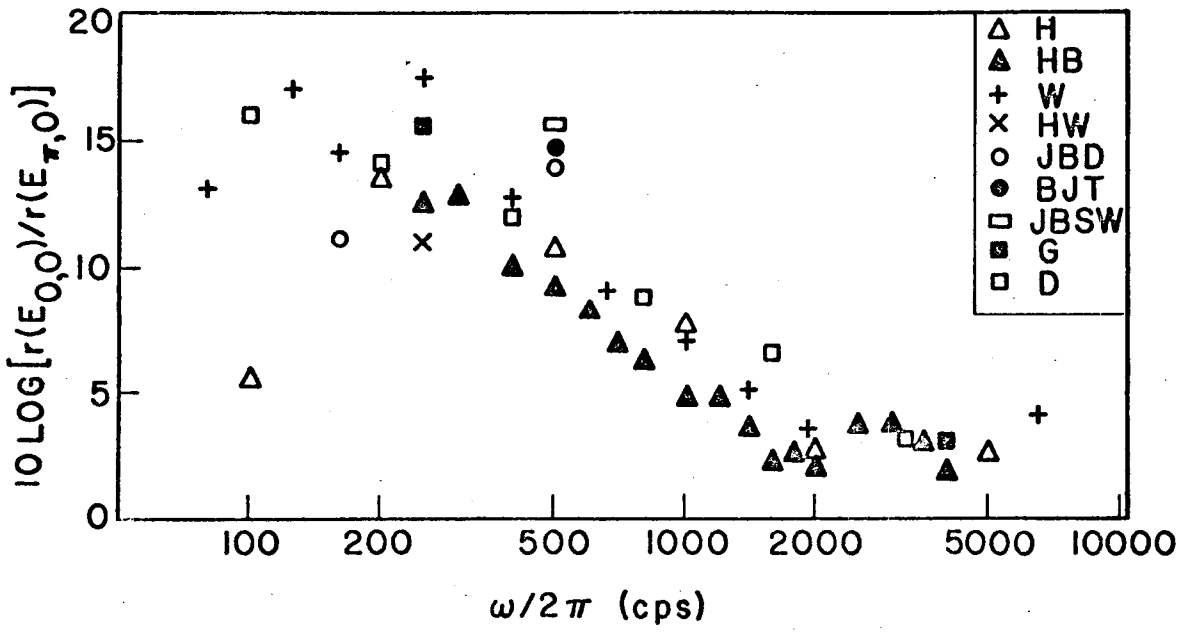
20
15
10
5
0
10 LOG [r(E0,0)/r(Eπ,0)]
100
200
500
1000
2000
5000
10000
ω/2π (cps)
H
HB
W
HW
JBD
BJT
JBSW
G
D

Fig. 5

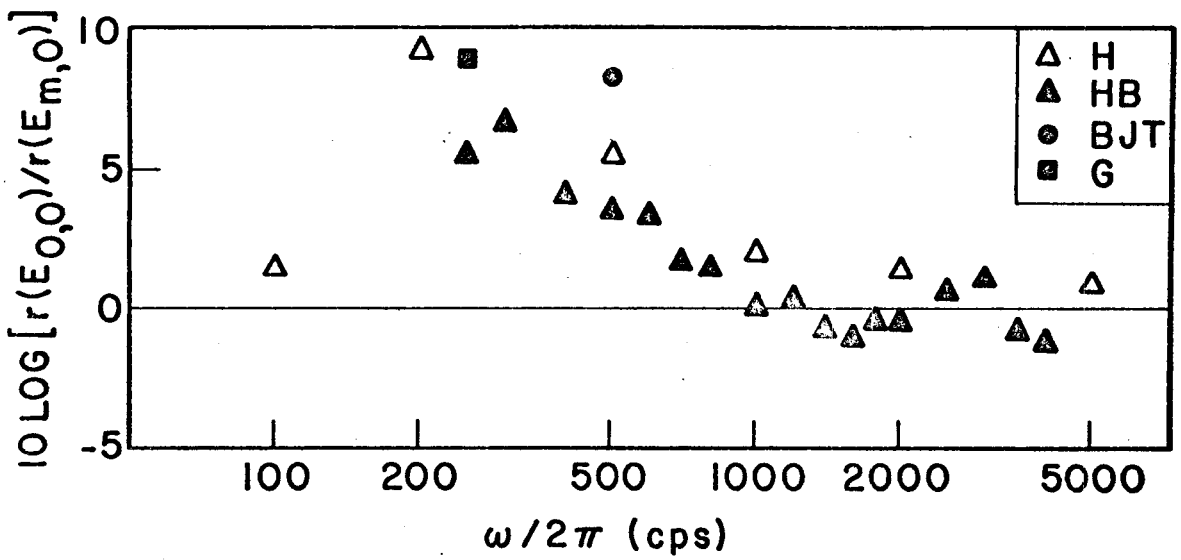
10 LOG [r(E0,0)/r(Em,0)]
10
5
0
-5
100
200
500
1000
2000
5000
ω/2π (cps)
H
HB
BJT
G

Fig. 6

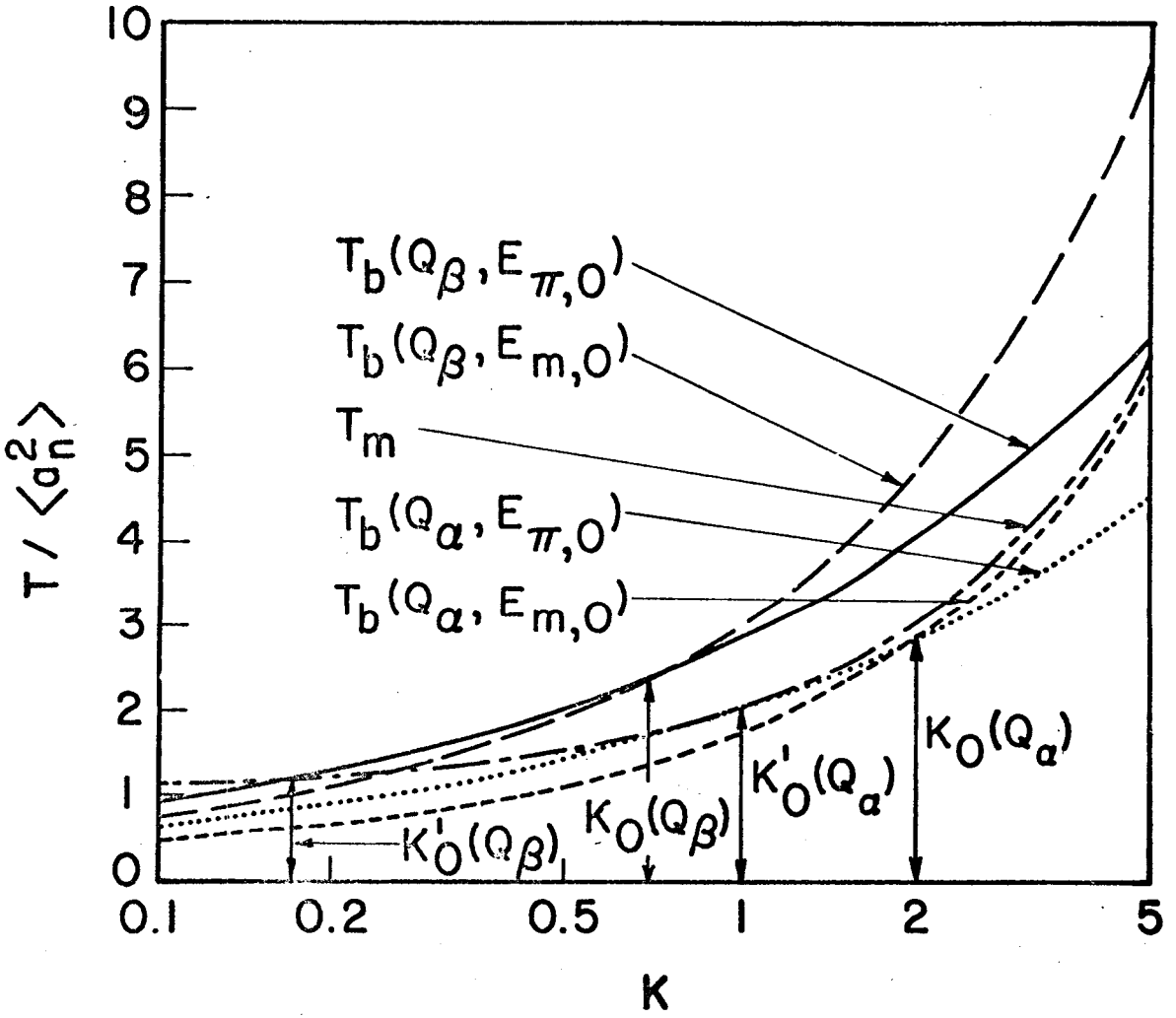
10
9
8
7
6
5
4
3
2
1
0
T / ⟨a_n^2⟩
T_b(Q_β, E_π,0)
T_b(Q_β, E_m,0)
T_m
T_b(Q_α, E_π,0)
T_b(Q_α, E_m,0)
K'_0(Q_β)
K_0(Q_β)
K'_0(Q_α)
K_0(Q_α)
0.1
0.2
0.5
1
2
5
K

Fig. 7